International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

# Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)

Herliana Noviar[1], Suryarini Widodo[2]
[1,2]Master of Information Systems Management Departement, Business Information System, Gunadarma University, Depok, West Java, Indonesia-16424
Email Address: [1]herliana1511@gmail.com, [2]srini@staff.gunadarma.ac.id

***Abstract*****—** *One important aspect that must be considered when creating a mobile application is usability. Usability is the extent to which a system, product, or service can be used by certain users to achieve certain goals with effectiveness, efficiency, and satisfaction in a certain context of use. Application usability assessment is very important when it is related to user satisfaction both in terms of assessing the interface and user experience. The application should also be easy to use so that users can get the information and services they need and continue to be loyal to using the application. Two well-known mobile language lesson applications are Duolingo and Cake. Based on ISO 9241-11, this research compares the Duolingo and Cake applications on mobile. The research instrument used to conduct this research was a questionnaire. Based on ISO 9241-11, the survey consists of 21 questions arranged based on four variables: effectiveness, efficiency, satisfaction and usability. The results of the questionnaire were analyzed using the Partial Least Square (PLS) method with the SmartPLS 4 tool which produced analysis results that the variables effectiveness, efficiency, satisfaction in the Duolingo and Cake mobile applications both had a positive direct and significant influence to usability.*

***Keywords*****—** *Usability Analysis, Mobile Application, Duolingo, Cake, ISO 9241-11, PLS.*

## I. INTRODUCTION

Technology is developing rapidly in all aspects of life in this digital era. In various countries there are new technologies that can help human life. One of the latest technologies is the presence of smartphones. As the use of smartphones continues to grow, of course this is a great opportunity to create mobile-based applications where all features can be accessed through applications that can be used on smartphones. Currently, the trend of using mobile-based applications is experiencing very rapid growth. One potential area in creating mobile applications is applications for learning foreign languages online.

Before technological advances, someone who wanted to learn a foreign language usually took a foreign language course which required a lot of time and energy. In this digital era, someone can learn a foreign language anywhere and anytime without being hindered by distance and time. Many language lesson applications have been developed by software developers, some of the famous ones are Duolingo and Cake.

With intense competition in the learning lesson application market, Duolingo and Cake can compete with excellence and steal the hearts of foreign language learning application users. Both have successfully reached the market in Indonesia and the world. It can be proven that these two applications have a large number of users throughout the world.

In general, mobile applications have a complex interface with various layers of menus and modules, so the application needs to be designed and developed to be more attractive and user friendly so that it can be accepted by its users. The usability aspect is the key to the success of an application and a condition for user acceptance so that they can obtain the information and services they need so that they continue to loyally use the application.

Measuring the usability level of an application can be done using a standard. Of the several standards available, ISO 9241-11 can be used as a reference because it provides a framework for measuring usability. As per ISO 9241-11 (1998), usability refers to the degree to which a product can be effectively and efficiently utilized by certain users to accomplish predetermined goals and attain user satisfaction within a given context. Usability assessment is very important when it is related to user satisfaction, both in terms of interface assessment and functionality. The variable criteria for usability according to Standard ISO 9241-11 are effectiveness, efficiency, and satisfaction (Anonymus, 1998). There are several methods to determine the relationship between these three variables and usability, one method that can be used is the Partial Least Square (PLS) method. PLS is an alternative method for structural equation modeling, namely to simultaneously test the relationship between latent constructs and many indicators. PLS is often referred to as soft modeling because it eliminates assumptions, such as the assumption that the data does not have to have a multivariate normal distribution, does not assume that the data must be on a certain scale of measurement, and the sample size does not have to be large (Hair et al, 2014).

Researchers are interested in conducting research on the Duolingo and Cake mobile applications because both have good ratings and a high number of users (observation results on Google Play, 2023). As of October 2023, Duolingo has a rating of 4.7 out of 5.0 on Google Play. On the other hand, Cake has a rating of 4.8 out of 5.0 on Google Play. Therefore, researchers conducted an analysis of the Duolingo and Cake learning lesson applications to find out to what extent the applications are usable for their users because the two applications have similarities in terms of service provision and



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

to determine the effectiveness, efficiency and satisfaction of their users.

## II. Literature Review

### *2.1. Analysis*

The definition of analysis is an activity in studying and evaluating a form of problem or case that occurs (Santi, 2020). According to Komaruddin (2001), analysis is a thinking activity to decompose a whole into components so that you can recognize the signs of the components, their relationship to each other and the function of each in an integrated whole. From the opinions above, it can be concluded that analysis is an activity in investigating or studying an event (unit) into smaller units to obtain a fact.

### *2.2. Usability*

According to Wahyuningrum (2021), usability is the end user's perception of how someone can effectively, efficiently and satisfactorily complete tasks when using an application. The definition of usability according to Nielsen (1993), namely the context of overall system acceptance. Usability ensures the product is easy to learn, effective to use, and enjoyable from the user's perspective (Sharp et al., 2002). Usability in dictionary is defined as "capability of being used" (Bevan et al, 2015). Accordingly, ISO 9241-11 (Bevan et al, 2015) defined usability as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”. According to ISO 9241-11 (1998), to measure usability, it is necessary to measure effectiveness, efficiency and satisfaction to achieve goals.

### *2.3. Mobile Application*

Mobile applications or also known as mobile applications (mobile devices) are a type of application software that is intended to run on mobile phones, for example smartphones or tablet PCs. Mobile applications often serve to provide users with services similar to those on a PC (Putra et al, 2023).

### *2.4. Duolingo*

Luis von Ahn and Severin Hacker are the creators of the free language learning program Duolingo. Apart from being available in the web version, this application is also available in the Android and iOS versions. As of 2022, the application provides more than 100 courses teaching more than 40 languages and all of them are free (Duolingo, 2022).

### *2.5. Cake*

Cake is an English and Korean learning application that has been used by 100 million users worldwide. Cake's learning method uses free video (vlogs, movies, TV shows) and audio content that is updated daily (Google Play, 2023). This application is available in Android and IOS versions.

### *2.6. ISO 9241-11*

The goal of designing and evaluating systems, products and services for usability is to enable users to achieve goals effectively, efficiently and satisfactorily by considering the context of use. The ISO 9241-11 (2018) standard highlights that usability is contingent upon the particular conditions under which a system, product, or service is utilized and provides an explanation of how usability might be interpreted in connection to user performance and pleasure. ISO 9241-11 (2018) explains how to interpret each component in the definition of usability: “the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”. It is measured by the extent to which the intended use goal is achieved, the resources that must be expended to achieve the intended goal, and the extent to which the user considers use of the product acceptable.

### *2.7. Likert Scale*

The Likert scale is a tool used to gauge an individual's or a group's attitudes, beliefs, and perceptions on social issues. This social phenomenon—henceforth referred to as the research variable—has been precisely identified by the researcher for the purposes of the study. The variables to be measured are converted into variable indicators using a Likert Scale. Then, these indicators serve as a basis for the compilation of instrument items, which may take the kind of questions or statements (Sugiyono, 2022).

For the purposes of quantitative analysis, the answers can be given a score, as in the table 2.1.

TABLE 2.1. Assessment Weight

| Number | Interval Weights | Score |
|---|---|---|
| 1 | Agree / Always / Very Positive | 5 |
| 2 | Agree / Often / Positive | 4 |
| 3 | Undecided / Sometimes / Neutral | 3 |
| 4 | Disagree / Almost Never / Negative | 2 |
| 5 | Strongly Disagree / Never | 1 |

Source: (Sugiyono, 2022)

### *2.8. Partial Least Square (PLS)*

Partial Least Square (PLS) was first developed by Herman Wold (1982). PLS is a powerful analysis method because it does not assume the data must be on a certain measurement scale and also involves a relatively small sample size (recommended minimum ranges from 30 to 500). The PLS approach is distribution free, which means that data cannot have a certain distribution, it can be nominal, categorical, ordinal, interval and ratio. In its development, the basic PLS model was completed by Herman Wold in 1977, which was then further developed by Lohmoller in 1984 and 1989, and then developed by Chin in 1996 (Ghozali, 2011).

According to Jogiyanto and Abdillah (2009), PLS can be used for variant-based Structural Equation Modeling (SEM) analysis which can simultaneously test measurement models as well as test structural models. While the structural model is used to assess causality (testing hypotheses with prediction models), the measurement model is used to test validity and reliability.

### *2.9. Preparation of Path Analysis and Measurement Model*

Based on the basic theory of PLS, a flow diagram (path diagram) of causal relationships between constructs and indicators can be created. Path Analysis was developed as a



Herliana Noviar and Suryarini Widodo, “Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS),” *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

method for studying the influence (effect) of independent variables on dependent variables.

The measurement model or often called the outer model shows how the manifest variables (observed variables or indicators) represent the latent variables (unobserved variables or constructs) to be measured. The measurement model shows how each indicator block is related to its latent variable (Ghozali, 2021). In measuring latent variables (constructs), there are two types of measurement models which can be identified as follows:

*a) Construct with reflective indicators*

Constructs with reflexive indicators assume that the covariance between model measurements is explained by variance that is a manifestation of the construct domain. The direction of the indicators is from construct to indicator.

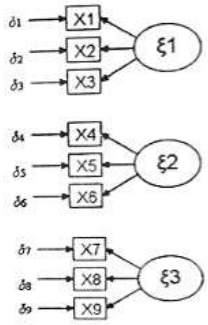


Fig. 2.1 Constructs with Reflective Indicators
Source: (Ghozali, 2021)

*b) Construct with formative indicators*

Constructs with formative indicators assume that each indicator defines or explains the characteristics of the construct domain. The direction of the indicators is from indicator to construct. Measurement error is directed at the construct, not at the indicator, so testing the validity and reliability of the construct is not necessary.

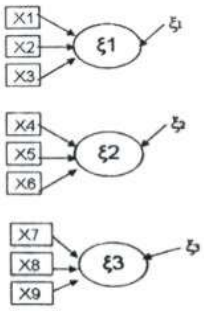


Fig. 2.2 Constructs with Formative Indicators
Source: (Ghozali, 2021)

### *2.10. Testing the Validity of the Research Model*

Validity testing with a reflexive indicator model is used to identify that unobserved variables (latent variables) can be measured using each observed variable construct (manifest variable) through Confirmatory Factor Analysis (CFA) or better known as confirmatory factor analysis. In general, before carrying out structural model analysis, researchers must first carry out a measurement model to test the validity and reliability of the indicators forming the latent construct by conducting confirmatory factor analysis (CFA) (Ghozali, 2021). If the factor loading value of each construct is more than 0.7, then it is declared valid (but some research states that validity values above 0.5 or 0.6 are still acceptable for measurements to be developed), or in other words, unobserved variables can be measured using each observed variable construct (Imam et al., 2012).

The results of the test will get a loading factor value for each construct. Then, using SmartPLS 4 software, the Convergent Validity, Discriminant Validity and Avarage Variance Extracted (AVE) values will be calculated in accordance with the standard validity values mentioned above for the main questionnaire and Microsoft Excel for the pretest questionnaire.

Convergent validity of the measurement model with reflexive indicators can be seen from the correlation between the item/indicator scores and the construct scores. Individual indicators are considered reliable if they have a correlation value above 0.70. Nonetheless, loadings between 0.50 and 0.60 are still appropriate at this point of research scale development (Ghozali, 2021).

The model has sufficient discriminant validity if the AVE root for each construct is greater than the correlation between the construct and other constructs. A good model if the AVE value of each construct is greater than 0.50. A recommended AVE value of greater than 0.50 indicates that at least half of the indicator's variance may be explained (Ghozali, 2021).

### *2.11. Testing the Reliability of the Research Model*

Reliability is the reliability to which a questionnaire, test, observation or any measurement is carried out. Reliability is a procedure that produces stability or consistency of scores from an assessment. This reliability is the result of existing assessments from the source. This is said to be reliable if the results of each measurement carried out are consistent in the same data (Primananda Putra, 2014). In the reliability analysis, it is hoped that you will be able to find out that the questions for each questionnaire are related to each other. Apart from that, it is also to get scaling of the overall measurement (Cronbach Alpha). Reliability is also useful for selecting questionnaires, if there is a questionnaire that has problems with the questions, you can delete it.

To test the reliability of the pretest questionnaire, the reliability test used was the Cronbach Alpha reliability test using Microsoft Excel. The statistical Cronbach Alpha coefficient is very well used as an indicator of the reliability of a questionnaire which shows that the subject provides responses with the same pattern during a certain period.

Reliability testing for the main questionnaire with the reflective indicator model will be carried out by calculating the Composite Reliability and Cronbach Alpha of each item in a variable using SmartPLS 4 software. If the Cronbach Alpha and Composite Reliability scores are greater than 0.70, the construct is deemed dependable (Ghozali, 2021).



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

## III. RESEARCH METHOD

In order to ensure that the research is conducted with proper direction, the approach that will be employed in the process will now be carefully outlined.

### *3.1. Research Steps*

Based on Figure 3.1, the activities carried out in this research are as follows:

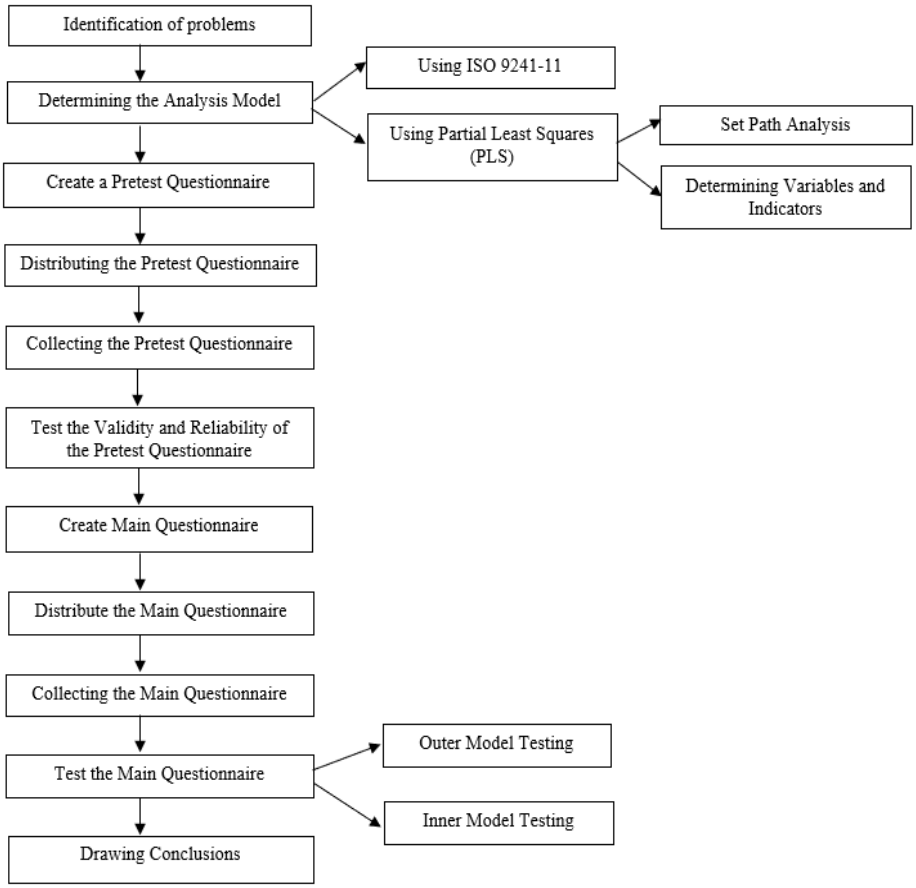


Fig. 3.1 Research Method

- Stage 1: Identify the problem that will be researched. In this research, the topic raised is the mobile learning lesson application Duolingo and Cake.
- Stage 2: The analysis model that is deemed suitable and accurate for this research is Partial Least Square (PLS) using a usability approach based on ISO 9241-11, because the main objective of this research is to test the hypothesis of the effectiveness, efficiency and satisfaction variables on the usability of mobile applications Duolingo and Cake. Next, the author makes a path analysis and determines the variables and indicators.
- Stage 3: Pretest questionnaire is a feasibility test of the questionnaire based on the opinions of several first respondents and is carried out before distributing the main questionnaire. The question points included in the pretest questionnaire are indicators of the variables that were created in Stage 2. The measurement procedure is that the first 30 respondents who used the Duolingo application and the first 30 respondents who used the Cake application were asked to express their agreement based on the perceptions of each respondent.
- Stage 4: Distribution of the pretest questionnaire was carried out to several users of the Duolingo and Cake mobile applications who were deemed capable of providing answers regarding usability and could represent the answers of all mobile application users of the Duolingo and Cake mobile applications.
- Stage 5: Before processing, the data is collected and saved in CVS (comma delimited) file format or in Excel file format to be tested for validity and reliability using the functions in Excel.
- Stage 6: Pretest questionnaire data is processed to test the validity and reliability of each question. If the questions are declared valid and reliable, then they can be used for the main questionnaire.
- Stage 7: The questions contained in the main questionnaire are questions that are valid and reliable from the results of the questionnaire's pretest validity and reliability tests.
- Stage 8: Distribution of the main questionnaire was carried out to several users of the Duolingo and Cake mobile applications who were deemed capable of providing answers regarding usability and could represent the



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

answers of all mobile application users of the Duolingo and Cake mobile applications.

- Stage 9: Before processing, the data is collected and saved in CVS (comma delimited) file format or in Excel file format because SmartPLS 4 can read files input in that format.
- Stage 10: Data obtained from respondents will then be processed using the Partial Least Square (PLS) method with the SmartPLS 4 tool. Analysis will be carried out based on hypothesis testing based on the variables measured.
- Stage 11: At this stage, results will be obtained regarding the influence of effectiveness, efficiency and satisfaction on usability as well as an explanation of design recommendations which can be used as a reference for developing this application.

### *3.2. Explanation Research Steps*

#### *3.2.1. Identification of Problems*

Researchers are interested in conducting research on the Duolingo and Cake mobile applications because both have good ratings and a high number of users (observation results on Google Play, 2023). As of October 2023, Duolingo has a rating of 4.7 out of 5.0 on Google Play. On the other hand, Cake has a rating of 4.8 out of 5.0 on Google Play. Therefore, researchers conducted an analysis of the Duolingo and Cake learning lesson applications to find out to what extent the applications are usable for their users because the two applications have similarities in terms of providing services.

#### *3.2.2. Determining the Analysis Model*

The analysis model that is deemed suitable for this research is Partial Least Square (PLS) using a usability approach based on ISO 9241-11. ISO 9241-11 can be used as a reference because it provides a framework for measuring usability. According to ISO 9241-11 (1998), to measure usability, it is necessary to measure effectiveness, efficiency and satisfaction to achieve goals. Partial Least Square (PLS) can be used to measure ISO 9241-11 variables, namely effectiveness, efficiency and satisfaction, on usability by evaluating the measurement model (outer model) and structural model (inner model). The measurement model (outer model) is used to test validity and reliability, while the structural model (inner model) is used to test causality (testing hypotheses with prediction models).

##### *3.2.2.1. Determining Variables and Indicators*

Research variables are basically anything in any form that is determined by the researcher to be studied so that information about it is obtained and then conclusions are drawn (Sugiyono, 2021). The variables used in this research are based on ISO 9241-11. The variable criteria for usability according to ISO 9241-11 are effectiveness, efficiency and satisfaction. Research variables and indicators can be seen in table 3.1.

TABLE 3.1. Research Variables and Indicators

| Variable | Indicator | Code |
|---|---|---|
| Effectiveness | The menu in the Duolingo/Cake application is easy to understand | EFE01 |
| | The information displayed in the Duolingo/Cake application is in accordance with the user's wishes | EFE02 |
| | The layout in the Duolingo/Cake application is correct | EFE03 |
| | Task completion is easy to understand | EFE04 |
| | Has a responsive help feature if you experience difficulties in completing assignments or questions | EFE05 |
| | Availability of reminders/notifications to continue learning | EFE06 |
| Efficiency | The Duolingo/Cake application encourages users to be enthusiastic about learning | EFI01 |
| | The menu in the Duolingo/Cake application is responsive when clicked | EFI02 |
| | The Duolingo/Cake application displays a fast response after the user answers a question | EFI03 |
| | Presentation of learning content that is easy to learn | EFI04 |
| | Quick Duolingo/Cake app error/bug fix | EFI05 |
| Satisfaction | Users are satisfied using the Duolingo/Cake application | SAT01 |
| | Users are interested in using this application again | SAT02 |
| | The learning content presented in the Duolingo/Cake application is updated | SAT03 |
| | The Duolingo/Cake application is convenient to use | SAT04 |
| | Organized learning content according to user level | SAT05 |
| Usability | The Duolingo/Cake app provides benefits to users | USA01 |
| | The Duolingo/Cake application makes it easier for users to recognize foreign language learning | USA02 |
| | The Duolingo/Cake app is easy to operate | USA03 |
| | The menu display in the Duolingo/Cake application is appropriate | USA04 |
| | The learning content on the Duolingo/Cake application is in accordance with user needs | USA05 |

#### *3.2.3. Create a Pretest Questionnaire*

The main instrument in this research is a questionnaire. Measurement of each indicator in the variable is carried out using a Likert Scale.

The pretest questionnaire is a feasibility test of the questionnaire based on the opinions of several first respondents and is carried out before distributing the main questionnaire. The purpose of the pretest questionnaire is to find out how much the respondent understands the questions given in the questionnaire.

There are five answer choices available, namely: Strongly Disagree (STS), Disagree (TS), Neutral (N), Agree (S), and Strongly Agree (SS). For the purposes of quantitative analysis, answers can be given a score as in table 3.2.

TABLE 3.2. Weight of Questionnaire Answers

| Assessment Answers | Abbreviation | Weight Value |
|---|---|---|
| Strongly Disagree | STS | 1 |
| Disagree | TS | 2 |
| Neutral | N | 3 |
| Agree | S | 4 |
| Strongly Agree | SS | 5 |

Source: (Sugiyono, 2022)

Sampling in this study used simple random sampling, namely sampling that gives each element of the population the same opportunity to be taken.



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

*3.2.4. Distributing the Pretest Questionnaire*

Sampling was carried out by randomly distributing questionnaires. For the pretest questionnaire, testing will be carried out on the first 30 respondents and after obtaining the results of validity and reliability testing, it will continue with distributing the main questionnaire by taking valid and reliable questions obtained from the previous pretest questionnaire. Distribution of the questionnaire was carried out using Google Form.

*3.2.5. Collecting the Pretest Questionnaire*

At this stage, the results of the pretest questionnaire are collected and recorded. All pretest questionnaire results were quantitatively calculated for each respondent and then processed and analyzed using Excel.

*3.2.6. Test the Validity and Reliability of the Pretest Questionnaire*

To test the validity of each question item for the pretest questionnaire, Microsoft Excel was used for the calculations. Calculating the validity of an instrument can use the Product Moment correlation formula or also known as Pearson Correlation. Where for decision making, if r is positive and r calculated ≥ r table then the question item is valid. Conversely, if r is negative and r count < r table then the question item is invalid

To test the reliability of the pretest questionnaire, the reliability test used was the Cronbach Alpha reliability test using Microsoft Excel. The statistical Cronbach Alpha coefficient is very well used as an indicator of the reliability of a questionnaire which shows that the subject provides responses with the same pattern during a certain period. In the reliability analysis, it is hoped that you will be able to find out that the questions for each questionnaire are related to each other. Reliability is also useful for selecting questionnaires, if there is a questionnaire that has problems with the questions, you can delete it.

*3.2.7. Create Main Questionnaire*

The question points (indicators) contained in the main questionnaire are questions that are valid and reliable from the results of the questionnaire's pretest validity and reliability tests. Measurement of each indicator in the variable is carried out using a Likert scale.

Sampling in this study used simple random sampling, namely sampling that gives each element of the population the same opportunity to be taken. The sample size for the questionnaire in this study was adjusted to the analysis model used, namely Partial Least Square (PLS) between 30 - 500 samples. Apart from that, determining the sample size also uses the Charlez Gervitsz method which is adapted to the PLS analysis model. This method is guided by the population size in sampling with the division as follows (Azwar, 2012):

- Minimum sample of 30, if < 30 is considered too small to describe the conclusions drawn
- Population > 500 people, taken as 10%
- Population > 5000 people, taken 100 – 500

*3.2.8. Distribute the Main Questionnaire*

The distribution of the main questionnaire (main questionnaire) was carried out to 250 user respondents, namely 125 respondents for the Duolingo mobile application and 125 respondents for the Cake mobile application according to the Partial Least Square (PLS) analysis model. Distribution of the questionnaire was carried out using Google Form.

*3.2.9. Collect Main Questionnaire*

At this stage, the results of the main questionnaire are collected and recorded. All results of the main questionnaire were quantitatively calculated for each respondent and then data processing and analysis was carried out using the SmartPLS 4 tool.

*3.2.10. Test the Main Questionnaire*

*3.2.10.1. Measurement Model Testing (Outer Model)*

In general, before carrying out structural model analysis (inner model), researchers must first carry out a measurement model (outer model) to test the validity and reliability of the indicators forming the latent construct by conducting confirmatory factor analysis (CFA) (Ghozali, 2021).

*3.2.10.1.1 Validity Test*

To test the validity of the main questionnaire, it is carried out based on the Partial Least Square (PLS) analysis model and using the SmartPLS 4 tool which will directly test the validity using Confirmatory Factor Analysis (CFA). The results of the test will get a loading factor value for each construct. Then, using SmartPLS 4 software, the Convergent Validity, Discriminant Validity and Avarage Variance Extracted (AVE) values will be calculated.

*3.2.10.1.2 Reliability Test*

To test the reliability of the main questionnaire, it will be carried out based on the Partial Least Square (PLS) analysis model and using the SmartPLS 4 tool. Reliability testing using the reflective indicator model will be carried out by calculating the Composite Reliability and Cronbach Alpha for each item in a variable using the SmartPLS 4 software.

*3.2.10.2. Measurement Structural Model Testing (Inner Model)*

Predicting the link between latent variables is the goal of structural model evaluation, also known as inner model evaluation. The structural model (inner model) shows the strength of estimates between latent variables or constructs. The inner model is evaluated by looking at the percentage of variance explained, namely by looking at the R-Square value for the endogenous latent construct (Ghozali, 2021). This test consists of 2 stages, namely the R Square Determinant Coefficient ($R^2$) test, which is a test that calculates how much the independent (exogenous) latent variable explains the variance of the dependent (endogenous) latent variable and the hypothesis test which is a test of the research model hypothesis.

*3.2.10.2.1 Hypothesis Testing*

To see the significance to determine the influence between variables, SmartPLS 4 uses the bootstrap resampling method. The significance value is tested using the t-test to determine whether or not the proposed hypothesis is accepted. The T-Statistics test is calculated by comparing the T-count value with the T-Table. Significant means that the research hypothesis that has been proven in the sample can be applied to the population (Sugiyono, 2022). Significant means that a research hypothesis that has been proven in a sample can be

generalized to the entire population from which the sample was taken.

The hypothesis proposed is:

H0: The Efficiency/Effectiveness/Satisfaction variables in the Duolingo and Cake mobile applications have a direct and significant positive effect on Usability.

H1: The Efficiency/Effectiveness/Satisfaction variables in the Duolingo and Cake mobile applications have a direct and insignificant positive effect on Usability.

*3.2.10.2.2 Hypothesis Test Results*

Hypothesis testing is carried out by looking at the magnitude of the T-Statistics value using a significance level of 95% (α = 0.05). The one tail T-Table value with a significance level of 95% is 1.65. The T-Statistics test is calculated by comparing the T-Count value with the T-Table.

The significance level (α) shows the probability or chance of error that the researcher determines in making a decision to reject or support the null hypothesis. In this research, the researcher took a significance level of 0.05 or 5%, meaning that the confidence level of this research is 95% or with the following conditions:

- Significance ≥ 0.05 then H0 is accepted
- Significance < 0.05 then H0 is rejected.

## IV. Research Result and Discussion

### *4.1. Respondent Characteristics*

The number of respondents used in this research was 250 respondents, namely 125 users of the Duolingo application service and 125 users of the Cake application service. The following are the characteristics of all respondents based on the respondent's gender and age:

#### *4.1.1. Duolingo Application User Respondents*

The following are the characteristics of 125 respondents who use the Duolingo application:

##### *4.1.1.1. Duolingo Application Respondent Data by Gender*

Based on the table 4.1, it can be seen that the majority of respondents were female, namely 80 respondents, while the remaining 45 respondents were male.

TABLE 4.1. Distribution of Duolingo Application Respondents by Gender

| Gender | Total (n) | Percentage (%) |
|---|---|---|
| Man | 45 | 36% |
| Woman | 80 | 64% |

##### *4.1.1.2. Duolingo Application Respondent Data by Age*

TABLE 4.2. Distribution of Duolingo Application Respondents by Age

| Age | Total (n) | Percentage (%) |
|---|---|---|
| < 20 years | 31 | 25% |
| 21 – 25 years | 50 | 40% |
| 26 – 30 years | 11 | 9% |
| 31 – 35 years | 14 | 11% |
| > 35 years | 19 | 15% |

Based on the table 4.2 above, it can be seen that most of the respondents were aged 21-25 years.

##### *4.1.1.3. Duolingo Application Respondent Data by Job*

Based on the table 4.3, it can be seen that the majority of respondents work as students.

TABLE 4.3. Distribution of Duolingo Application Respondents by Job

| Job | Total (n) | Percentage (%) |
|---|---|---|
| Student | 69 | 55% |
| Private Employees/Civil Servants | 24 | 19% |
| Housewife | 13 | 10% |
| Entrepreneur | 19 | 15% |

#### *4.1.2. Cake Application User Respondents*

The following are the characteristics of 125 respondents who use the Cake application:

##### *4.1.2.1. Cake Application Respondent Data based on Gender*

TABLE 4.4. Distribution of Cake Application Respondents by Gender

| Gender | Total (n) | Percentage (%) |
|---|---|---|
| Man | 51 | 41% |
| Woman | 74 | 59% |

Based on the table above, it can be seen that the majority of respondents were female, namely 74 respondents, while the remaining 51 respondents were male.

##### *4.1.2.2. Cake Application Respondent Data by Age*

TABLE 4.5. Distribution of Cake Application Respondents by Age

| Age | Total (n) | Percentage (%) |
|---|---|---|
| < 20 years | 24 | 19% |
| 21 – 25 years | 38 | 30% |
| 26 – 30 years | 20 | 16% |
| 31 – 35 years | 14 | 11% |
| > 35 years | 29 | 23% |

Based on the table above 4.5, it can be seen that most of the respondents were aged 21-25 years.

##### *4.1.2.3. Cake Application Respondent Data based on Job*

TABLE 4.6. Distribution of Cake Application Respondents by Job

| Job | Total (n) | Percentage (%) |
|---|---|---|
| Student | 45 | 36% |
| Private Employees/Civil Servants | 34 | 27% |
| Housewife | 24 | 19% |
| Entrepreneur | 22 | 18% |

Based on the table above 4.6, it can be seen that the majority of respondents work as students.

### *4.2. Questionnaire*

Below is data on the answers to the results of the respondent's questionnaire. The results of the questionnaire are divided into two, namely the results on the pretest questionnaire (initial test of the questionnaire) and the results on the main questionnaire.

#### *4.2.1. Pretest Questionnaire*

Pretest questionnaire or initial test of the questionnaire is a feasibility test of the questionnaire based on the opinions of several first respondents and is carried out before distributing the main questionnaire. The purpose of the initial questionnaire test is to find out how much the respondent understands the questions given in the questionnaire.

The pretest questionnaire consisted of 21 questions and was administered to 30 respondents to the Duolingo application and 30 respondents to the Cake application. Then the data is processed to test the validity and reliability of each



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

question. If the questions are declared valid and reliable, then the questionnaire can be distributed to a total of 250 respondents, namely 125 respondents who use the Duolingo application and 125 respondents who use the Cake application.

*4.2.1.1. Questionnaire Pretest Validity Test*

Research that produces good quality data needs to test the validity of the pretest questionnaire before the main questionnaire is given. This validity test is used to provide an understanding that the questionnaire being measured is able to provide true value. The steps for testing the validity of the pretest questionnaire (using the Product Moment correlation formula using Microsoft Excel). With the number of respondents in the pretest totaling 30 people and a significance value of 5%, the r table which is used as a reference is 0.361. If the calculated r correlation value is greater than or equal to r table (sig. 5%), then the question is valid. However, if the calculated r correlation value is smaller than table r (sig. 5%), then the question is invalid.

The following are the results of the validity test on the pretest questionnaire for each criterion in the Duolingo and Cake applications:

a). Test the Pretest Validity of the Duolingo Application Questionnaire

TABLE 4.7. Pretest Validity Test of the Duolingo Application User Questionnaire

| Variable | Question | Corrected Total Item Correlation | Conclusion |
|---|---|---|---|
| Effectiveness | EFE01 | 0,641 | Valid |
| | EFE02 | 0,758 | Valid |
| | EFE03 | 0,709 | Valid |
| | EFE04 | 0,608 | Valid |
| | EFE05 | 0,726 | Valid |
| | EFE06 | 0,467 | Valid |
| Efficiency | EFI01 | 0,435 | Valid |
| | EFI02 | 0,755 | Valid |
| | EFI03 | 0,682 | Valid |
| | EFI04 | 0,633 | Valid |
| | EFI05 | 0,636 | Valid |
| Satisfaction | SAT01 | 0,829 | Valid |
| | SAT02 | 0,741 | Valid |
| | SAT03 | 0,690 | Valid |
| | SAT04 | 0,768 | Valid |
| | SAT05 | 0,670 | Valid |
| Usability | USA01 | 0,775 | Valid |
| | USA02 | 0,805 | Valid |
| | USA03 | 0,708 | Valid |
| | USA04 | 0,629 | Valid |
| | USA05 | 0,837 | Valid |

Based on table 4.7, it can be concluded that all question items from each variable, namely Effectiveness, Efficiency, Satisfaction and Usability have valid values.

b). Test the Pretest Validity of the Cake Application Questionnaire

TABLE 4.8. Pretest Validity Test of the Cake Application User Questionnaire

| Variable | Question | Corrected Total Item Correlation | Conclusion |
|---|---|---|---|
| Effectiveness | EFE01 | 0,829 | Valid |
| | EFE02 | 0,794 | Valid |
| | EFE03 | 0,797 | Valid |
| | EFE04 | 0,816 | Valid |
| | EFE05 | 0,895 | Valid |
| | EFE06 | 0,758 | Valid |
| Efficiency | EFI01 | 0,837 | Valid |
| | EFI02 | 0,873 | Valid |
| | EFI03 | 0,904 | Valid |
| | EFI04 | 0,887 | Valid |
| | EFI05 | 0,860 | Valid |
| Satisfaction | SAT01 | 0,903 | Valid |
| | SAT02 | 0,918 | Valid |
| | SAT03 | 0,836 | Valid |
| | SAT04 | 0,868 | Valid |
| | SAT05 | 0,832 | Valid |
| Usability | USA01 | 0,877 | Valid |
| | USA02 | 0,882 | Valid |
| | USA03 | 0,903 | Valid |
| | USA04 | 0,920 | Valid |
| | USA05 | 0,919 | Valid |

Table 4.8 show that it can be concluded that all question items from each variable, namely Effectiveness, Efficiency, Satisfaction and Usability have valid values.

*4.2.1.2. Questionnaire Pretest Reliability Test*

The aim of the reliability test is to determine the connectedness of each question on the questionnaire. To test the reliability of the pretest questionnaire, the method used was the Cronbach Alpha reliability test using Microsoft Excel. The statistical Cronbach Alpha coefficient is very well used as an indicator of the reliability of a questionnaire which shows that the subject provides responses with the same pattern during a certain period. If the Cronbach Alpha value is $\geq 0.70$, then the question item is reliable, whereas if the Cronbach Alpha value is $\leq 0.70$, then the question item is not reliable. The following are the results of the reliability test on the pretest questionnaire for each criterion in the Duolingo and Cake applications:

a). Duolingo Application Questionnaire Pretest Reliability Test

Based on table 4.9, it can be concluded that the four variables, namely Effectiveness, Efficiency, Satisfaction, and Usability have a high level of reliability.

TABLE 4.9. Pretest Reliability Test of the Duolingo Application User Questionnaire

| Variable | Number of Questions | Invalid Question | Cronbach's Alpha Value | Explanation |
|---|---|---|---|---|
| Effectiveness | 6 | 0 | 0,777 | High Reliability |
| Efficiency | 5 | 0 | 0,720 | High Reliability |
| Satisfaction | 5 | 0 | 0,852 | Very High Reliability |
| Usability | 5 | 0 | 0,884 | Very High Reliability |

b). Cake Application Questionnaire Pretest Reliability Test

TABLE 4.10. Pretest Reliability Test for Cake Application User Questionnaire

| Variable | Number of Questions | Invalid Question | Cronbach's Alpha Value | Explanation |
|---|---|---|---|---|
| Effectiveness | 6 | 0 | 0,928 | Very High Reliability |
| Efficiency | 5 | 0 | 0,929 | Very High Reliability |



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

| | | | | |
|---|---|---|---|---|
| Satisfaction | 5 | 0 | 0,949 | Very High Reliability |
| Usability | 5 | 0 | 0,955 | Very High Reliability |

Based on table 4.10, it can be concluded that the four variables, namely Effectiveness, Efficiency, Satisfaction, and Usability have a very high level of reliability (0.8 > a < 1).

*4.2.1.3. Recapitulation of Pretest Questionnaire*

From testing the validity and reliability of the pretest questionnaire, it can be concluded that the pretest questionnaire is valid and reliable so that all the 21 pretest questionnaire questions can be used for the main questionnaire.

*4.2.2. Main Questionnaire*

The main questionnaire is the main questionnaire that is distributed to all target respondents for data collection and processing. The main questionnaire contains all pretest questionnaire questions that have been declared valid and reliable and have their own indicators for processing and creating Path Diagram models. After testing the validity and reliability of the pretest questionnaire, it can be concluded that all questions can be included in the main questionnaire. Thus, the number of main questionnaire questions is the same as the number of pretest questionnaire questions, namely 21 questions. The main questionnaire was distributed to a total of 250 respondents, namely 125 respondents using the Duolingo application and 125 respondents using the Cake application.

*4.2.2.1. Duolingo Application*

*4.2.2.1.1. Measurement Model Testing (Outer Model)*

This model defines how each indicator is related to its latent variable, or it could be said that the outer model can specify the relationship between the latent variable and its indicators. Tests carried out on the outer model used the Confirmatory Factor Analysis (CFA) technique.

In testing the data, researchers distributed questionnaires to 125 Duolingo user respondents. The following is an overview of the data that has been modeled in a path diagram and will be processed with SmartPLS 4.0.

In Figure 4.1, all independent variables and dependent variables have been entered into the Path Diagram. The independent variable effectiveness with 6 indicators (EFE01 - EFE06), the independent variable efficiency with 5 indicators (EFI01 – EFI05), the independent variable satisfaction with 5 indicators (SAT01 – SAT05), and the dependent variable usability with its 5 indicators (USA01 – USA05). These four variables will be tested for validity and reliability before testing the hypothesis. The following is an overview of the model after the values for each indicator are input and processed using the PLS Algorithm.

In Figure 4.2, the variable factor loading values and indicators for each construct are obtained which will be used for further testing.

*4.2.2.1.1.1. Validity Test*

Based on the PLS method, testing the validity of reflexive indicators is carried out in 2 stages.

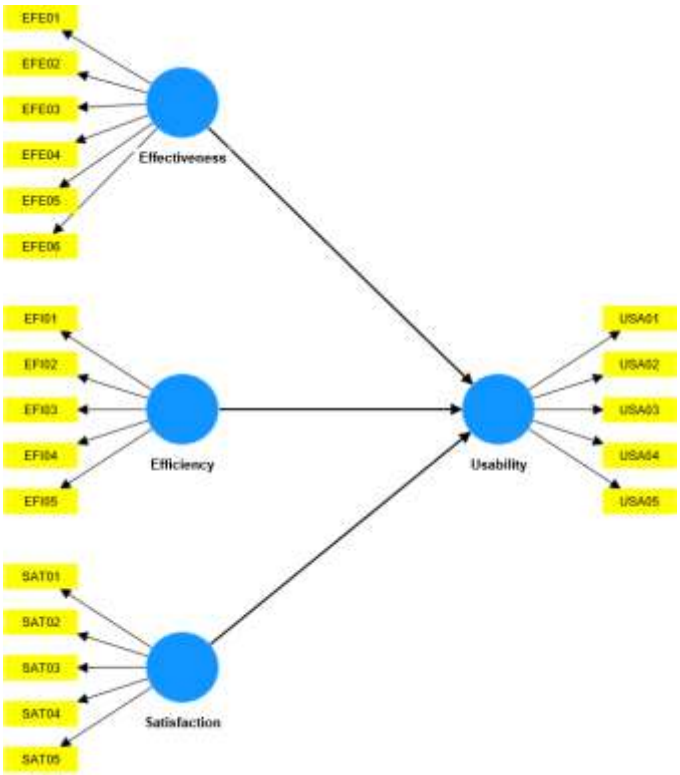


Fig. 4.1 Usability Construct Model for the Duolingo Application Using SmartPLS 4.0

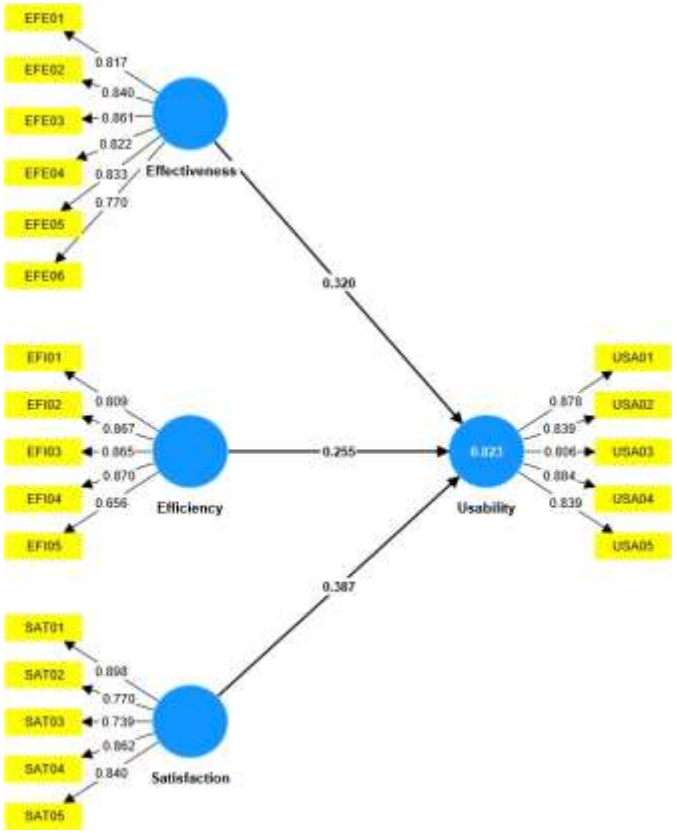


Fig. 4.2 Value Model Between Constructs Usability Research Model for the Duolingo Application Using SmartPLS 4.0



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

The first stage is convergent validity testing, namely validity testing based on the loading factor value of each construct, and the next stage is discriminant validity testing, namely validity testing based on a comparison of the magnitude of the correlation values between constructs.

a) Convergent Validity

The first stage of validity testing is used to identify that unobserved variables can be measured using each observed variable construct through Confirmatory Factor Analysis (CFA) or what is usually called factor analysis. Indicators are said to meet convergent validity if the loading factor value is above 0.7, however for research in the early stages of development a loading value of 0.5 to 0.6 is considered sufficient (Ghozali, 2021). The results of validity testing are shown in the table 4.11.

TABLE 4.11. Loading Factor Values for Each Variable and Construct of the Duolingo Application Usability Research Model

| Variable | Indicator Code | Outer Loading Value |
|---|---|---|
| Effectiveness | EFE01 | 0,817 |
| | EFE02 | 0,840 |
| | EFE03 | 0,861 |
| | EFE04 | 0,822 |
| | EFE05 | 0,833 |
| | EFE06 | 0,770 |
| Efficiency | EFI01 | 0,809 |
| | EFI02 | 0,867 |
| | EFI03 | 0,865 |
| | EFI04 | 0,870 |
| | EFI05 | 0,656 |
| Satisfaction | SAT01 | 0,898 |
| | SAT02 | 0,770 |
| | SAT03 | 0,739 |
| | SAT04 | 0,862 |
| | SAT05 | 0,840 |
| Usability | USA01 | 0,878 |
| | USA02 | 0,839 |
| | USA03 | 0,806 |
| | USA04 | 0,884 |
| | USA05 | 0,839 |

From table 4.11 it can be seen that all loading factor values for each indicator are above 0.5. This proves that all the indicators used in this research are valid or have met convergent validity. The results in the table above are the results of outer loading for each indicator belonging to each latent variable obtained from data processing using SmartPLS.

b) Discriminant Validity

The second stage of validity testing is discriminant validity testing. The measurement's cross loading value with the construct and the Average Variance Extracted (AVE) value serve as the foundation for this test. Cross loading factor to determine whether the latent variable has adequate discriminant, namely by comparing the correlation of the indicator with the latent variable, it must be greater than the correlation between the indicator and other latent variables (Sofyan Yamin, 2009). If the correlation value of the construct with the measurement item is greater than the correlation value with other constructs, then this shows that the latent construct predicts the indicators in their block better than the indicators in other blocks, and this is what is called each construct has unidimensionality so it is said that the construct has high discriminant validity (Ghozali, 2021). The following are the discriminant validity results from the cross loading values between the indicators and their respective constructs:

TABLE 4.12. Cross Loading Values for Each Variable and Construct of the Duolingo Application Usability Research Model

| | Effectiveness | Efficiency | Satisfaction | Usability |
|---|---|---|---|---|
| EFE01 | 0,817 | 0,707 | 0,674 | 0,759 |
| EFE02 | 0,840 | 0,732 | 0,674 | 0,736 |
| EFE03 | 0,861 | 0,707 | 0,692 | 0,702 |
| EFE04 | 0,822 | 0,732 | 0,724 | 0,709 |
| EFE05 | 0,833 | 0,769 | 0,685 | 0,727 |
| EFE06 | 0,770 | 0,669 | 0,634 | 0,621 |
| EFI01 | 0,693 | 0,809 | 0,654 | 0,652 |
| EFI02 | 0,776 | 0,867 | 0,706 | 0,735 |
| EFI03 | 0,704 | 0,865 | 0,623 | 0,724 |
| EFI04 | 0,769 | 0,870 | 0,693 | 0,782 |
| EFI05 | 0,622 | 0,656 | 0,579 | 0,523 |
| SAT01 | 0,709 | 0,726 | 0,898 | 0,760 |
| SAT02 | 0,617 | 0,535 | 0,770 | 0,571 |
| SAT03 | 0,619 | 0,613 | 0,739 | 0,610 |
| SAT04 | 0,695 | 0,676 | 0,862 | 0,698 |
| SAT05 | 0,743 | 0,704 | 0,840 | 0,835 |
| USA01 | 0,756 | 0,716 | 0,749 | 0,878 |
| USA02 | 0,746 | 0,707 | 0,681 | 0,839 |
| USA03 | 0,661 | 0,666 | 0,663 | 0,806 |
| USA04 | 0,789 | 0,789 | 0,763 | 0,884 |
| USA05 | 0,708 | 0,699 | 0,770 | 0,839 |

From table 4.12, it can be seen that the correlation value of the construct with its indicators is greater than the correlation value with other constructs. Thus it can be concluded that all latent constructs show good discriminant validity because they can predict indicators in their block better than indicators in other blocks.

Next, discriminant validity testing is carried out by looking at the Average Varience Extracted (AVE) value. The AVE value is good if it has a value greater than 0.50 (Ghozali, 2021). The following are the values from the AVE table.

Table 4.13 shows the AVE value of the research model. It can be seen from the table that the AVE value for all research variables has a value above 0.5 so that the AVE value for discriminant validity testing is sufficient for further testing.

TABLE 4.13. AVE from the Duolingo Application Usability Research Model

| Variable | AVE Value |
|---|---|
| Effectiveness | 0,679 |
| Efficiency | 0,668 |
| Satisfaction | 0,679 |
| Usability | 0,722 |

Thus, the discriminant validity test has been fulfilled as well as the convergent validity test so that it can be concluded that the Duolingo Application usability research model is valid.

*4.2.2.1.1.2. Reliability Test*

Based on the PLS method, the reliability of the reflective indicators in this study is determined from the Composite Reliability and Cronbach Alpha values for each block of first order indicators in the reflective construct. A construct is declared reliable if the Composite Reliability and Cronbach Alpha values are above 0.70 (Ghozali, 2021).

a) Composite Reliability



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

Another test to evaluate the outer model is to look at the reliability of the latent variable construct which is measured by two types of measures, namely Composite Reliability and Cronbach Alpha from the block of indicators that measure the construct. In the first stage, the construct is declared reliable if it has a Composite Reliability value above 0.70 (Imam Ghozali, 2021). The following are the output results from the outer model of Composite Reliability:

TABLE 4.14. Composite Reliability Values from the Duolingo Application Usability Research Model

| Variable | Composite Reliability |
|---|---|
| Effectiveness | 0,927 |
| Efficiency | 0,909 |
| Satisfaction | 0,913 |
| Usability | 0,928 |

Table 4.14 above is a table of Composite Reliability values from the research model. The table shows that each variable has a composite reliability value above 0.7 with the lowest value being 0.909 from the Efficiency variable and the highest value being 0.928 from the Usability variable. From these results it can be concluded that the research model has met the value of Composite Reliability.

b) Cronbach Alpha

The next stage of reliability testing is testing the Cronbach Alpha value. A construct is declared reliable if it has a Cronbach Alpha value above 0.70 (Imam Ghozali, 2021). The following are the output results from the outer model from Cronbach Alpha:

TABLE 4.15. Cronbach Alpha Value from the Duolingo Application Usability Research Model

| Variable | Cronbach Alpha |
|---|---|
| Effectiveness | 0,905 |
| Efficiency | 0,873 |
| Satisfaction | 0,881 |
| Usability | 0,903 |

Table 4.15 above is a table of Cronbach Alpha values from the research model. The table shows that each variable has a Cronbach Alpha value above 0.7 with the lowest value being 0.873 from the Efficiency variable and the highest value being 0.905 from the Effectiveness variable. From these results it can be concluded that the research model has met the value of Cronbach Alpha.

*4.2.2.1.2. Structural Model Testing (Inner Model)*

This model is a specification of the relationship between latent variables, also called inner relations. This test is a test of the type and magnitude of the influence of the independent latent variable on the dependent latent variable. This test consists of 2 stages, namely the R Square Determinant Coefficient ($R^2$) test, which is a test that calculates how much the independent latent variable explains the variance of the dependent latent variable and the hypothesis test which is a test of the research model hypothesis.

*4.2.2.1.2.1. Coefficient of Determination Test / R Square ($R^2$)*

Evaluation of the inner model is carried out by looking at the Determination Coefficient. The Coefficient of Determination aims to measure how far the model's ability to explain the variance of the dependent variable. The coefficient of determination value is between 0 and 1. The coefficient of determination ($R^2$) value is close to 1. The $R^2$ value explains how much the independent variable hypothesized in the equation is able to explain the dependent variable. Ghozali (2021) explains that the R-Squares ($R^2$) values are 0.75, 0.50 and 0.25, it can be concluded that the model is strong, moderate and weak.

If the coefficient of determination value is small or below or equal to 0.500 ($R^2 \leq 0.500$), it means that the ability of the independent variables to explain variations in the dependent variable is very limited. Meanwhile, if the coefficient of determination value is greater than 0.500 ($R^2 > 0.500$), it means that the ability of the independent variables provides almost all the information needed to predict variations in the dependent variable (Imam Ghozali, 2011). The following is $R^2$ on the construct based on the output from SmartPLS 4:

TABLE 4.16. R Square ($R^2$) Value from the Duolingo Application Usability Research Model using SmartPLS 4

| Variable | R Square ($R^2$) |
|---|---|
| Usability | 0,823 |

Based on table 4.16, it can be seen that the R Square value of the usability variable is 0.823, which shows that the contribution of the effectiveness, efficiency and satisfaction variables to usability is 82.3%. Meanwhile, the remaining 17.7% of the usability variance of the Duolingo application is influenced by other factors outside the effectiveness, efficiency and satisfaction variables. The R Square ($R^2$) value of 0.823 which exceeds 0.75 indicates that the model is in the strong category.

*4.2.2.1.2.2. Hypothesis Testing*

SmartPLS 4.0 uses a nonparametric test to determine the significance level of the path coefficient, where the T (T-Statistics) value produced by running the bootstrapping algorithm in SmartPLS 4.0 is used to determine whether or not the proposed hypothesis is accepted.

The way to test hypotheses using SmartPLS 4.0 is to look at the Original Sample Estimate, Standard Error, and i values. The value in the Original Sample Estimate shows the type of relationship between the independent variable and the dependent variable. If the Original Sample Estimate value is positive, then the relationship between the two variables is positive, which means that if there is an increase in the value of the independent variable, it will be followed by an increase in the value of the dependent variable. Vice versa (Ghozali, 2011).

Next, look at the T-Statistics value. This test can be done by comparing SmartPLS T-Statistics with T-Tables.

The significance level (α) shows the probability or chance of error that the researcher determines in making a decision to reject or support the null hypothesis. In this research, the researcher took a significance level of 0.05 or 5%, meaning that the confidence level of this research is 95% or with the following conditions:

- Significance ≥ 0.05 then H0 is accepted
- Significance < 0.05 then H0 is rejected



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

With degrees of freedom (df) of 121 (n - k = 121, n = 125, k = 4, k is the number of parameters/variables used in the research) and a significance level of 0.05, then if you look at the degrees of freedom (df) table below get a one tail t table value of 1.65. Therefore, if the T-Statistic is above or equal to 1.65 (T-Statistic ≥ 1.65), then the influence exerted by the independent variable on the dependent variable is significant. Meanwhile, if the T-Statistic is below 1.65 (T-Statistic < 1.65) then the influence given is not significant.

The following table 4.17 is a of Path Coefficients (Mean, STDEV, T-Values) resulting from SmartPLS 4.0 bootstrapping which is the basis for basic decision making:

TABLE 4.17. Path Coefficients (Mean, STDEV, T-Values) from the Duolingo Application Usability Research Model using SmartPLS 4.0

| Variable | Original sample (O) | Standard deviation (STDEV) | T statistics (O/STDEV) | Significance Level (1,65) |
|---|---|---|---|---|
| Effectiveness -> Usability | 0,320 | 0,105 | 3,061 | Significant |
| Efficiency -> Usability | 0,255 | 0,104 | 2,448 | Significant |
| Satisfaction -> Usability | 0,387 | 0,092 | 4,216 | Significant |

*4.2.2.1.2.3. Hypothesis Test Results*

From the structural model that has been formed, the coefficient of the relationship between the independent variable and the dependent variable is obtained. This coefficient is the coefficient of the relationship between latent variables and the contribution value of the manifest variables that form the latent variable. The relationships that occur in this structural model are the basis for evaluating the research hypothesis.

The significance level of each relationship between latent variables can be seen from the T-Statistics value which must be greater than the T-Table (1.65) for a significant effect and less than the table for an insignificant effect (α = 0.05). Significant means that the research hypothesis that has been proven in the sample can be applied to the population (Sugiyono, 2022). Significant means that a research hypothesis that has been proven in a sample can be generalized to the entire population from which the sample was taken.

The following is a test of the hypothesis H0 in this research:

1. *"Effectiveness of the Duolingo application has a direct and significant positive effect on usability"*

The table above shows that the relationship between Effectiveness and Usability is significant with a T-Statistic of 3.061 (T-Statistic > 1.65). The original sample estimate value is positive, namely 0.320, which indicates that the direction of the relationship between Effectiveness and Usability is positive. Thus, the hypothesis H0 in this study which states that "Effectiveness in the Duolingo application has a direct and significant positive effect on Usability" is accepted.

2. *"Efficiency in the Duolingo application has a direct and significant positive effect on usability."*

The table above shows that the relationship between Efficiency and Usability is significant with a T-Statistic of 2.448 (T-Statistic > 1.65). The original sample estimate value is positive, namely 0.255, which indicates that the direction of the relationship between Efficiency and Usability is positive. Thus, the hypothesis H0 in this study which states that "Efficiency in the Duolingo application has a direct and significant positive effect on Usability" is accepted.

3. *"Satisfaction with the Duolingo application has a direct and significant positive effect on usability"*

The table above shows that the relationship between Satisfaction and Usability is significant with a T-Statistic of 4.216 (T-Statistic > 1.65). The original sample estimate value is positive, namely 0.387, which indicates that the direction of the relationship between Satisfaction and Usability is positive. Thus, the hypothesis H0 in this study which states that "Satisfaction with the Duolingo application has a direct and significant positive effect on Usability" is accepted.

From the test results above, it is known that all hypotheses are accepted because they have a T-Statistic above 1.65. This shows that the aspects of Efficiency, Effectiveness, Satisfaction have a significant influence on Usability in the Duolingo application.

*4.2.2.2. Cake Application*

*4.2.2.2.1. Measurement Model Testing (Outer Model)*

This model defines how each indicator is related to its latent variable, or it could be said that the outer model can specify the relationship between the latent variable and its indicators. Tests carried out on the outer model used the Confirmatory Factor Analysis (CFA) technique.

In testing the data, researchers distributed questionnaires to 125 Cake user respondents. The following is an overview of the data that has been modeled in a path diagram and will be processed with SmartPLS 4.0.

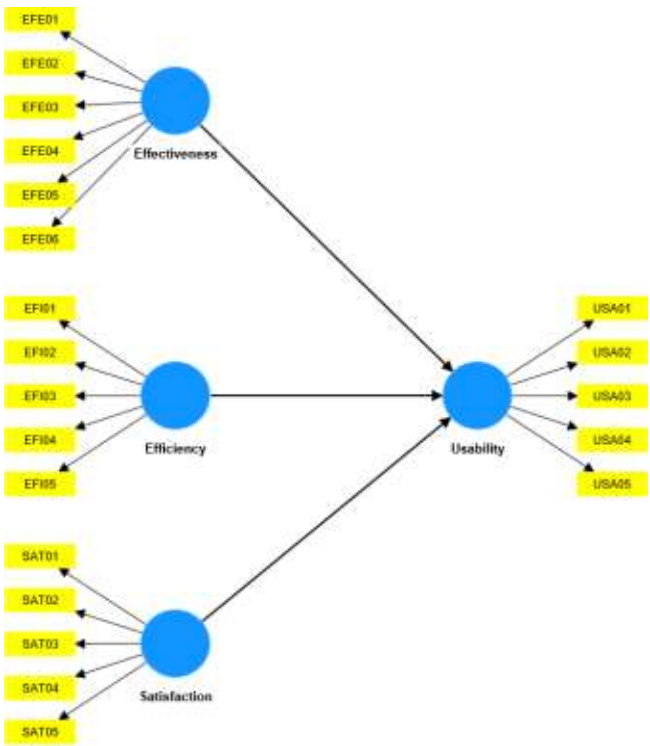


Fig. 4.3 Cake Application Usability Construct Model Using SmartPLS 4.0



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

In figure 4.3, all independent variables and dependent variables have been entered into the Path Diagram. The independent variable effectiveness with 6 indicators (EFE01 - EFE06), the independent variable efficiency with 5 indicators (EFI01 – EFI05), the independent variable satisfaction with 5 indicators (SAT01 – SAT05), and the dependent variable usability with its 5 indicators (USA01 – USA05). These four variables will be tested for validity and reliability before testing the hypothesis. The following is an overview of the model after the values for each indicator are input and processed using the PLS Algorithm.

Figure 4.4 show that the variable factor loading values and indicators for each construct are obtained which will be used for further testing.

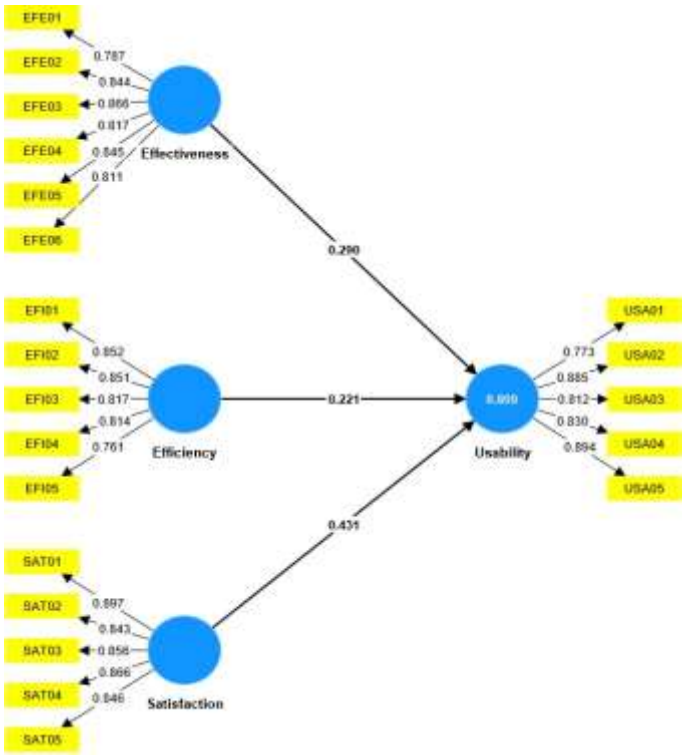


Fig. 4.4. Interconstruct Value Model Cake Application Usability Research Model Using SmartPLS 4.0

*4.2.2.2.1.1. Validity test*

Based on the PLS method, testing the validity of reflexive indicators is carried out in 2 stages. The first stage is convergent validity testing, namely validity testing based on the loading factor value of each construct, and the next stage is discriminant validity testing, namely validity testing based on a comparison of the magnitude of the correlation values between constructs.

a) Convergent Validity

The first stage of validity testing is used to identify that unobserved variables can be measured using each observed variable construct through Confirmatory Factor Analysis (CFA) or what is usually called factor analysis. Indicators are said to meet convergent validity if the loading factor value is above 0.7, however for research in the early stages of development a loading value of 0.5 to 0.6 is considered sufficient (Ghozali, 2021). The results of validity testing are shown in table 4.18.

TABLE 4.18. Loading Factor Values for Each Variable and Construct of the Cake Application Usability Research Model

| Variable | Indicator Code | Outer Loading Value |
|---|---|---|
| Effectiveness | EFE01 | 0,787 |
| | EFE02 | 0,844 |
| | EFE03 | 0,866 |
| | EFE04 | 0,817 |
| | EFE05 | 0,845 |
| | EFE06 | 0,811 |
| Efficiency | EFI01 | 0,852 |
| | EFI02 | 0,851 |
| | EFI03 | 0,817 |
| | EFI04 | 0,814 |
| | EFI05 | 0,761 |
| Satisfaction | SAT01 | 0,897 |
| | SAT02 | 0,843 |
| | SAT03 | 0,856 |
| | SAT04 | 0,866 |
| | SAT05 | 0,846 |
| Usability | USA01 | 0,773 |
| | USA02 | 0,885 |
| | USA03 | 0,812 |
| | USA04 | 0,830 |
| | USA05 | 0,894 |

From table 4.18 it can be seen that all loading factor values for each indicator are above 0.5. This proves that all the indicators used in this research are valid or have met convergent validity. The results in the table above are the results of outer loading for each indicator belonging to each latent variable obtained from data processing using SmartPLS.

b) Discriminant Validity

The second stage of validity testing is discriminant validity testing. The measurement's cross loading value with the construct and the Average Variance Extracted (AVE) value serve as the foundation for this test. If the correlation value of the construct with the measurement item is greater than the correlation value with other constructs, then this shows that the latent construct predicts the indicators in their block better than the indicators in other blocks, and this is what is called each construct has unidimensionality so it is said that the construct has high discriminant validity (Ghozali, 2021). The following table 4.19 show the discriminant validity results from the cross loading values between the indicators and their respective constructs.

TABLE 4.19. Cross Loading Values for Each Variable and Construct of the Cake Application Usability Research Model

| | Effectiveness | Efficiency | Satisfaction | Usability |
|---|---|---|---|---|
| EFE01 | 0,787 | 0,673 | 0,665 | 0,669 |
| EFE02 | 0,844 | 0,728 | 0,731 | 0,759 |
| EFE03 | 0,866 | 0,733 | 0,744 | 0,724 |
| EFE04 | 0,817 | 0,687 | 0,694 | 0,646 |
| EFE05 | 0,845 | 0,722 | 0,733 | 0,741 |
| EFE06 | 0,811 | 0,640 | 0,662 | 0,641 |
| EFI01 | 0,756 | 0,852 | 0,751 | 0,773 |
| EFI02 | 0,715 | 0,851 | 0,771 | 0,775 |
| EFI03 | 0,681 | 0,817 | 0,722 | 0,640 |
| EFI04 | 0,650 | 0,814 | 0,699 | 0,642 |
| EFI05 | 0,641 | 0,761 | 0,690 | 0,620 |
| SAT01 | 0,777 | 0,808 | 0,897 | 0,797 |
| SAT02 | 0,696 | 0,763 | 0,843 | 0,692 |



Herliana Noviar and Suryarini Widodo, “Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS),” *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

| | | | | |
|---|---|---|---|---|
| SAT03 | 0,710 | 0,779 | 0,856 | 0,746 |
| SAT04 | 0,725 | 0,745 | 0,866 | 0,771 |
| SAT05 | 0,758 | 0,732 | 0,846 | 0,756 |
| USA01 | 0,577 | 0,615 | 0,682 | 0,773 |
| USA02 | 0,753 | 0,785 | 0,732 | 0,885 |
| USA03 | 0,697 | 0,691 | 0,785 | 0,812 |
| USA04 | 0,748 | 0,728 | 0,711 | 0,830 |
| USA05 | 0,753 | 0,733 | 0,759 | 0,894 |

From table 4.19, it can be seen that the correlation value of the construct with its indicators is greater than the correlation value with other constructs. Thus, it can be concluded that all latent constructs show good discriminant validity because they can predict indicators in their block better than indicators in other blocks.

Next, discriminant validity testing is carried out by looking at the Average Varience Extracted (AVE) value. The AVE value is good if it has a value greater than 0.50 (Ghozali, 2021). The following are the values from the AVE table:

TABLE 4.20. AVE from the Cake Application Usability Research Model

| Variable | AVE Value |
|---|---|
| Effectiveness | 0,687 |
| Efficiency | 0,671 |
| Satisfaction | 0,743 |
| Usability | 0,706 |

Table 4.20 shows the AVE value of the research model. It can be seen from the table that the AVE value for all research variables has a value above 0.5 so that the AVE value for discriminant validity testing is sufficient for further testing.

Thus, the discriminant validity test has been fulfilled as well as the convergent validity test so that it can be concluded that the Cake Application usability research model is valid.

*4.2.2.2.1.2. Reliability Test*

Based on the PLS method, the reliability of the reflective indicators in this study is determined from the Composite Reliability and Cronbach Alpha values for each block of first order indicators in the reflective construct. A construct is declared reliable if the Composite Reliability and Cronbach Alpha values are above 0.70 (Ghozali, 2021).

a) Composite Reliability

Another test to evaluate the outer model is to look at the reliability of the latent variable construct which is measured by two types of measures, namely Composite Reliability and Cronbach Alpha from the indicator block that measures the construct. In the first stage, the construct is declared reliable if it has a Composite Reliability value above 0.70 (Imam Ghozali, 2021). Table 4.21 show the output results from the outer model of Composite Reliability:

TABLE 4.21. Composite Reliability Values from the Cake Application Usability Research Model

| Variable | Composite Reliability |
|---|---|
| Effectiveness | 0,929 |
| Efficiency | 0,911 |
| Satisfaction | 0,935 |
| Usability | 0,923 |

Table 4.21 is a table of Composite Reliability values from the research model. The table shows that each variable has a composite reliability value above 0.7 with the lowest value being 0.911 from the Efficiency variable and the highest value being 0.935 from the Satisfaction variable. Based on these findings, it can be said that the research model satisfies the composite reliability requirement.

b) Cronbach Alpha

The next stage of reliability testing is testing the Cronbach Alpha value. A construct is declared reliable if it has a Cronbach Alpha value above 0.70 (Imam Ghozali, 2021). Table 4.22 show the output results from the outer model from Cronbach Alpha:

TABLE 4.22. Cronbach Alpha Value from the Cake Application Usability Research Model

| Variable | Cronbach Alpha |
|---|---|
| Effectiveness | 0,909 |
| Efficiency | 0,878 |
| Satisfaction | 0,913 |
| Usability | 0,895 |

Table 4.22 above is a table of Cronbach Alpha values from the research model. The table shows that each variable has a Cronbach Alpha value above 0.7 with the lowest value being 0.878 from the Efficiency variable and the highest value being 0.913 from the Satisfaction variable. From these results it can be concluded that the research model has met the value of Cronbach Alpha.

*4.2.2.2.2. Structural Model Testing (Inner Model)*

This model is a specification of the relationship between latent variables, also called inner relations. This test is a test of the type and magnitude of the influence of the independent latent variable on the dependent latent variable. This test consists of 2 stages, namely the R Square Determinant Coefficient ($R^2$) test, which is a test that calculates how much the independent latent variable explains the variance of the dependent latent variable and the hypothesis test which is a test of the research model hypothesis.

*4.2.2.2.2.1. Coefficient of Determination Test / R Square ($R^2$)*

Evaluation of the inner model is carried out by looking at the Determination Coefficient. The Coefficient of Determination aims to measure how far the model's ability to explain the variance of the dependent variable. The coefficient of determination value is between 0 and 1. The coefficient of determination ($R^2$) value is close to 1. The $R^2$ value explains how much the independent variable hypothesized in the equation is able to explain the dependent variable. Ghozali (2021) explains that the R-Squares ($R^2$) values are 0.75, 0.50 and 0.25, it can be concluded that the model is strong, moderate and weak.

If the coefficient of determination value is small or below or equal to 0.500 ($R^2 \leq 0.500$), it means that the ability of the independent variables to explain variations in the dependent variable is very limited. Meanwhile, if the coefficient of determination value is greater than 0.500 ($R^2 > 0.500$), it means that the ability of the independent variables provides almost all the information needed to predict variations in the dependent variable (Imam Ghozali, 2011). The following is $R^2$ on the construct based on the output from SmartPLS 4:



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

TABLE 4.23. R Square ($R^2$) Value from the Cake Application Usability Research Model using SmartPLS 4

| Variable | R Square ($R^2$) |
|---|---|
| Usability | 0,809 |

Based on table 4.23 above, it can be seen that the R Square value of the usability variable is 0.809, which shows that the contribution of the effectiveness, efficiency and satisfaction variables to usability is 80.9%. Meanwhile, the remaining 19.1% of the variance in Cake Application usability is influenced by other factors outside the variables of effectiveness, efficiency and satisfaction. The R Square ($R^2$) value of 0.809 which exceeds 0.75 indicates that the model is in the strong category.

*4.2.2.2.2.2. Hypothesis Testing*

SmartPLS 4.0 uses a nonparametric test to determine the significance level of the path coefficient, where the T (T-Statistics) value produced by running the bootstrapping algorithm in SmartPLS 4.0 is used to determine whether or not the proposed hypothesis is accepted.

The way to test hypotheses using SmartPLS 4.0 is to look at the Original Sample Estimate, Standard Error, and i values. The value in the Original Sample Estimate shows the type of relationship between the independent variable and the dependent variable. If the Original Sample Estimate value is positive, then the relationship between the two variables is positive, which means that if there is an increase in the value of the independent variable, it will be followed by an increase in the value of the dependent variable. Vice versa (Ghozali, 2011).

Next, look at the T-Statistics value. According to Ghozali, the T-Statistics value is the value of the influence of each independent variable partially (separately) on the dependent variable. This test can be done by comparing SmartPLS T-Statistics with T-Tables.

The significance level (α) shows the probability or chance of error that the researcher determines in making a decision to reject or support the null hypothesis. In this research, the researcher took a significance level of 0.05 or 5%, meaning that the confidence level of this research is 95% or with the following conditions:

- Significance ≥ 0.05 then H0 is accepted
- Significance < 0.05 then H0 is rejected

With degrees of freedom (df) of 121 (n - k = 121, n = 125, k = 4, k is the number of parameters/variables used in the research) and a significance level of 0.05, then if you look at the degrees of freedom (df) table below get a one tail t table value of 1.65. Therefore, if the T-Statistic is above or equal to 1.65 (T-Statistic ≥ 1.65), then the influence exerted by the independent variable on the dependent variable is significant. Meanwhile, if the T-Statistic is below 1.65 (T-Statistic < 1.65) then the influence given is not significant.

Table 4.24 below show the Path Coefficients (Mean, STDEV, T-Values) resulting from SmartPLS 4.0 bootstrapping which is the basis for basic decision making.

*4.2.2.2.2.3. Hypothesis Test Results*

From the structural model that has been formed, the coefficient of the relationship between the independent variable and the dependent variable is obtained. This coefficient is the coefficient of the relationship between latent variables and the contribution value of the manifest variables that form the latent variable. The relationships that occur in this structural model are the basis for evaluating the research hypothesis.

TABLE 4.24. Path Coefficients (Mean, STDEV, T-Values) from the Cake Application Usability Research Model using SmartPLS 4.0

| Variable | Original sample (O) | Standard deviation (STDEV) | T statistics (O/STDEV) | Significance Level (1,65) |
|---|---|---|---|---|
| Effectiveness -> Usability | 0,290 | 0,093 | 3,121 | Significant |
| Efficiency -> Usability | 0,221 | 0,121 | 1,821 | Significant |
| Satisfaction -> Usability | 0,431 | 0,123 | 3,492 | Significant |

The significance level of each relationship between latent variables can be seen from the T-Statistics value which must be greater than the T-Table (1.65) for a significant effect and less than the table for an insignificant effect (α = 0.05). A significant influence is a convincing or meaningful influence, in research meaning that a hypothesis that has been proven in the sample can be applied to the population (Ghozali, 2011).

The following is a test of the hypothesis H0 in this research:

1. *"Effectiveness of the Cake application has a direct and significant positive effect on usability"*

The table above shows that the relationship between Effectiveness and Usability is significant with a T-Statistic of 3.121 (T-Statistic > 1.65). The original sample estimate value is positive, namely 0.290, which indicates that the direction of the relationship between Effectiveness and Usability is positive. Thus, the hypothesis H0 in this study which states that "Effectiveness in the Cake application has a direct and significant positive effect on Usability" is accepted.

2. *"Efficiency in the Cake application has a direct and significant positive effect on usability"*

The table above shows that the relationship between Efficiency and Usability is significant with a T-Statistic of 1.821 (T-Statistic > 1.65). The original sample estimate value is positive, namely 0.221, which indicates that the direction of the relationship between Efficiency and Usability is positive. Thus, the hypothesis H0 in this study which states that "Efficiency in the Cake application has a direct and significant positive effect on Usability" is accepted.

3. *"Satisfaction (Satisfaction) with the Cake application has a direct and significant positive effect on usability"*

The table above shows that the relationship between Satisfaction (Satisfaction) and Usability is significant with a T-Statistic of 3.492 (T-Statistic > 1.65). The original sample estimate value is positive, namely 0.431, which indicates that the direction of the relationship between Satisfaction and Usability is positive. Thus, the hypothesis H0 in this study which states that "Satisfaction (Satisfaction) with the Cake application has a direct and significant positive effect on Usability" is accepted.



Herliana Noviar and Suryarini Widodo, "Comparative Analysis of Usability in Duolingo and Cake Mobile Applications Based on ISO 9241-11 Using Partial Least Square (PLS)," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 2, pp. 1-16, 2024.

From the test results above, it is known that all hypotheses are accepted because they have a T-Statistic above 1.65. This shows that the aspects of Efficiency, Effectiveness, Satisfaction have a significant influence on Usability in the Cake application.

### *4.3. Comparison of Analysis Results*

Based on the analysis explained above, in this study a comparison was obtained of the extent of the influence of Effectiveness, Efficiency and Satisfaction of users of the Duolingo and Cake applications on Usability. The results of the usability analysis comparison of the Duolingo and Cake applications can be seen in Table 4.25:

TABLE 4.25. Comparison of Usability Analysis Results for the Duolingo and Cake Applications

| Variable | Duolingo Application | Cake Application |
|---|---|---|
| Effectiveness -> Usability | Direct and significant positive influence | Direct and significant positive influence |
| Efficiency -> Usability | Direct and significant positive influence | Direct and significant positive influence |
| Satisfaction -> Usability | Direct and significant positive influence | Direct and significant positive influence |

## V. CONCLUSION

Based on the analysis of tests using the SmartPLS 4.0 tools that have been carried out in this research, it can be concluded that:

1) This research shows that the Effectiveness of the Duolingo and Cake applications both have a direct and significant positive influence on Usability.
2) This research shows that Efficiency in the Duolingo and Cake applications both have a direct and significant positive influence on Usability.
3) This research shows that Satisfaction with the Duolingo and Cake applications both have a direct and significant positive influence on Usability.

## REFERENCES

[1] Anonymous (2018). "ISO 9241-11". URL: https://cdn.standards.iteh.ai/samples/16883/44acafdfd9a24edd9c66ed2f0e2a50e2/ISO-9241-11-1998.pdf [Accessed 20 October 2023].
[2] App Store (2023). "Cake - Belajar Bahasa Asing". URL: https://apps.apple.com/id/app/cake-belajar-bahasa-asing/id1350420987?l=id [Accessed 19 October 2023].
[3] Cake (2018). "About Cake app". URL: https://www.cakecorp.com/view/en/index.html [Accessed 19 October 2023].
[4] Anonymous. (2020). "Cake ". URL: https://cake.day/ [Accessed 19 October 2023].
[5] Duolingo (2019). "Executive Management". URL: https://investors.duolingo.com/governance [Accessed 19 October 2023].
[6] Duolingo (2022). "2022 Duolingo Language Report". URL: https://blog.duolingo.com/2022-duolingo-language-report/ [Accessed 19 October 2023].
[7] Duolingo (2023). "Duolingo: Bahasa Inggris untuk penutur Bahasa Indonesia". URL: https://www.duolingo.com/enroll/en/id/Pelajari-Bahasa%20Inggris [Accessed 19 October 2023].
[8] Duolingo (2023). "Duolingo: Tentang Kami". URL: https://www.duolingo.com/info [Accessed 24 October 2023].
[9] Ghozali, I. (2021). "Partial Least Square: Konsep, Teknik dan Aplikasi Menggunakan Program SmartPLS 3.2.9 untuk Penelitian Empiris". Edisi 3. Semarang: Badan Penerbit Universitas Diponegoro.
[10] Ghozali, I. and Kusumadewi, K. (2023). "Partial Least Square: Konsep, Teknik dan Aplikasi Menggunakan Program SmartPLS 4.0 untuk Penelitian Empiris". Edisi 1. Semarang: Yoga Pratama.
[11] Google Play (2023). "Duolingo: Belajar Bahasa". URL: https://play.google.com/store/apps/details?id=com.duolingo&hl=id&gl=US [Accessed 19 October 2023].
[12] Google Play (2023). "Cake - Belajar Bahasa Asing". URL: https://play.google.com/store/apps/details?id=me.mycake&hl=id&gl=US [Accessed 19 October 2023].
[13] Anonymous. (2023). "ISO: Global Standarts for Trusted Goods and Services". URL: https://www.iso.org/ [Accessed 24 October 2023].
[14] Putra, Y., Dawis, A., Novi, Natsir, F., Fitria, Widhiyanti, A., Hasan, F., Somantri, Maniah. (2023). "Pengantar Aplikasi Mobile". Sukabumi: Haura Utama.
[15] Santi, I. (2020). "Analisa Perancangan Sistem". Pekalongan: Penerbit NEM – Anggota IKAPI.
[16] Sugiyono. (2021). "Statistika Untuk Penelitian". Bandung: Alfabeta.
[17] Sugiyono. (2022). "Metode Penelitian Kuantitatif". Bandung: Alfabeta.
[18] Nova, Y., Yunitasari, T. (2021). "Usability Analysis on ISO 9241-11 Based Bibit and Bareksa Mutual Software Application Using Partial Least Square (PLS)". International Research Journal of Advanced Engineering and Science, Volume 6, Issue 2, pp. 45-48.
[19] Yassien, Eman. and Raja'a Masa'deh. (2017). "The Impact of ERP System's Usability on Enterprise Resource Planning Project Implementation Success via the Mediating Role of User Satisfaction". Journal of Management Research, Vol. 9, Issue. 3.
[20] Bevan, J. Carter, S. & Harker, A. (2015). "ISO 9241-11 Revised: What have we Learnt about Usability Since 1998?". In: M. Kurosu (Ed.), Human-Computer Interaction: Design and Evaluation, Lecture Notes in Computer Science 9169, Springer, pp. 143-151. URL: https://doi.org/10.1007/978-3-319-20901-2_13
[21] Bevan, N., Carter, J., Earthy, J., Geis, T., & Harker, S. (2016). "New ISO Standards for Usability, Usability Reports and Usability Measures". In International Conference on Human-Computer Interaction, pp. 268-278. URL: https://doi.org/10.1007/978-3-319-39510-4_25